



СИСТЕМНЫЙ АНАЛИЗ И УПРАВЛЕНИЕ

УДК 004.021; 519.254

О РАЗРАБОТКЕ ПАКЕТА КОМПЬЮТЕРНОЙ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ ВЫБОРА КОРРЕЛЯЦИОННО-РЕГРЕССИОННЫХ МОДЕЛЕЙ АНАЛИЗА И ПРОГНОЗИРОВАНИЯ ЭМПИРИЧЕСКИХ ДАННЫХ

ON THE ELABORATION OF A PACKAGE OF COMPUTER DECISION SUPPORT FOR THE CHOICE OF CORRELATION AND REGRESSION MODELS OF ANALYSIS AND FORECAST OF EMPIRICAL DATA

Н.П. Путивцева, Т.В. Зайцева, С.В. Игрунова, Е.В. Нестерова, О.П. Пусная
N.P. Putivzeva, T.V. Zaitseva, S.V. Igrunova, E.V. Nesterova, O.P. Pusnaya

Белгородский государственный национальный исследовательский университет, Россия, 308015, Белгород, ул. Победы, 85
Belgorod State National Research University, 85 Pobeda St, Belgorod, 308015, Russia

e-mail: putivzeva@bsu.edu.ru, zaitseva@bsu.edu.ru, igrunova@bsu.edu.ru, nesterova@bsu.edu.ru, pusnaya@bsu.edu.ru

Аннотация. В данной статье описан процесс разработки пакета компьютерной поддержки принятия решений для выбора корреляционно-регрессионных моделей. В результате анализа литературных источников был сделан вывод, что в существующих статистических и эконометрических пакетах отсутствуют возможности выдачи рекомендаций пользователю по целесообразности использования той или иной регрессионной модели и объяснения выбора оптимальной регрессионной модели, для конкретных данных.

Resume. This article describes the process of elaboration of a package of computer decision support for the choice of correlation and regression models. According to analysis of the literature, it was concluded that the existing statistical and econometric packages don't allow issuing recommendations to the user on whether to use a particular regression model and don't explain the choice of the optimal regression model, for specific data.

Ключевые слова: поддержка принятия решений, корреляционно-регрессионный анализ, прогнозирование, выбор оптимальной модели, компьютерная поддержка выбора.

Keywords: decision support, correlation and regression analysis, forecasting, selection of the optimal model, computer support of the process of choice.

Введение

Развитие экономики, усложнение экономических процессов и повышение требований к принимаемым управленческим решениям в области макро и микроэкономики потребовало более тщательного и объективного анализа реально протекающих процессов на основе привлечения современных математических и статистических методов.

Потребность в средствах статистического анализа эконометрических данных в различных областях деятельности, особенно в науке, очень велика, что и послужило причиной развития рынка компьютерных программ для статистической обработки эконометрических данных. На современном этапе невозможно представить эконометрическое исследование без применения информационных технологий. В настоящее время исследователю доступно большое количество разнообразных программных продуктов, которые могут быть использованы для решения задач построения регрессионных моделей, описывающих зависимость между эмпирическими данными. Наиболее популярными программными пакетами являются EVies, SPSS, STATA, STATISTICA, S-PLUS, Stadia, STATGRAPHICS, SYSTAT, Minitab.

В результате анализа литературных источников был сделан вывод, что существующие статистические и эконометрические пакеты являются достаточно громоздкими и дорогими. Зачастую они обладают избыточным функционалом, не используемым для учебных расчетов, но значительно усложняющим интерфейс. Бесплатные пакеты представляют собой усеченные расчетные модели. Но самым большим недостатком данных пакетов является отсутствие в представленных пакетах возмож-

ности выдачи рекомендаций пользователю по целесообразности использования той или иной регрессионной модели и о причинах выбора программой оптимальной регрессионной модели, описывающей конкретные эмпирические данные.

В связи с вышесказанным разработка пакета компьютерной поддержки принятия решений для выбора корреляционно-регрессионных моделей представляется актуальной.

Программа представляет собой аналитическую информационную систему для расчета экономических моделей, анализа рассчитанных показателей и выбора на основе анализа наиболее подходящей модели для каждого конкретного эмпирических данных.

Специализированный пакет построения и выбора корреляционно-регрессионных моделей предназначен для решения следующих задач:

- Построение основных моделей парной (линейной и нелинейной) и множественной регрессии;
- Выбор наиболее подходящей модели для описания эмпирических данных (представлено на рисунке 1);
- Оценка параметров;
- Оценка статистической значимости модели и ее параметров;
- Интерпретация результатов, выводы и пояснения системы.

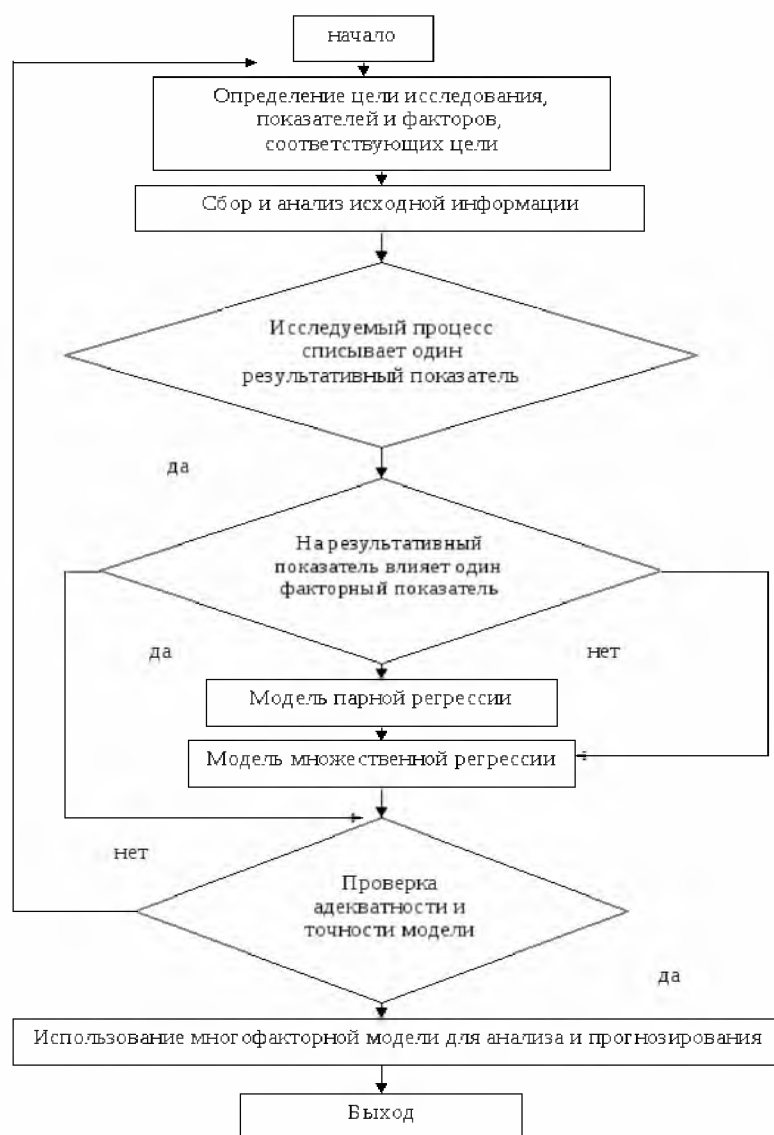


Рис.1. Алгоритм построения регрессионной модели
Fig. 1. An algorithm of construction of the regression model

Выходной информацией системы является:

1. Информация о типе выбранной модели с объяснением выбора наиболее подходящей модели.



2. Информация о параметрах уравнения.
3. Информация об оценке качества полученной модели.
4. Диаграммы.
5. Таблицы с расчетным значением зависимой переменной, точечным и интервальным прогнозом значений зависимой переменной.

Парная регрессия

- линейная $\hat{y}_x = a + b \cdot x$;
- равносторонняя гипербола $\hat{y}_x = a + b/x$;
- степенная $\hat{y}_x = a \cdot x^b$;
- показательная $\hat{y}_x = a \cdot b^x$;
- экспоненциальная $\hat{y}_x = e^{a+b \cdot x}$.

Множественная регрессия

$y = \hat{f}(x_1, x_2, \dots, x_m)$, где y – зависимая переменная, x_i – независимые, переменные.

Разработанное программное обеспечение предназначено для автоматизированной обработки эмпирической экономической информации и поддержки принятия решений выбора наиболее подходящей модели регрессии для описания имеющихся данных [1].

Выбор модели осуществляется на основе сравнения величин

$$\text{остаточной дисперсии } \sigma_{\text{ост}}^2 = \frac{1}{n} \sum (y - \hat{y}_x)^2,$$

$$\text{средней ошибки аппроксимации } \bar{A} = \frac{1}{n} \sum \left| \frac{y - \hat{y}_x}{y} \right| \cdot 100\%,$$

$$\text{коэффициента или индекса детерминации } r_{xy}^2 = 1 - \frac{\sigma_{\text{ост}}^2}{\sigma_y^2}, \quad \rho_{xy}^2 = 1 - \frac{\sigma_{\text{ост}}^2}{\sigma_y^2},$$

рассчитанных при разных моделях.

Если при выборе между использованием линейной и нелинейной моделей парной регрессии значения соответствующих показателей детерминации близки, система делает вывод о допустимости использования парной линейной регрессии, в противном случае сообщает о невозможности замены нелинейной регрессии линейной. Далее система определяет, к какому классу нелинейных регрессий относится наиболее подходящая для описания эмпирических данных функция. Если регрессия нелинейна по оцениваемым переменным, то выполняется линеаризация переменных, если же она нелинейна по оцениваемым параметрам, то система определяет, является ли данная нелинейная модель внутренне линейной или внутренне нелинейной. Если модель является внутренне линейной, система делает преобразование исходной нелинейной модели для приведения ее к линейному виду.

В том случае, если ни одна из полученных моделей парной регрессии не является значимой либо если величина возмущений сравнима с результатами, рассчитанными по построенным моделям, система предлагает построить множественную регрессию, введя дополнительные исходные данные значений того фактора, который целесообразно ввести в модель [4].

При построении моделей множественной регрессии система вначале определяет, будет ли модель множественной регрессии достаточно адекватно отражать зависимость между эмпирическими значениями, или же необходимо построить модели нелинейной множественной регрессии и на основе расчета и сравнения аналогичных показателей выбрать наиболее подходящую модель.

В случае адекватности линейной множественной регрессии система строит уравнение множественной регрессии в стандартизованном и натуральном масштабе переменных, выдает выводы о степени влияния каждого из факторов на зависимый показатель, о том, как ведет себя зависимый показатель при изменении каждого из факторов на единицу при неизменном значении остальных факторов. Система вычисляет критерии Фишера и Стьюдента, определяет статистическую значимость уравнения в целом и отдельных его параметров. На основе расчета и анализа частных критериев Фишера система выдает рекомендации о том, целесообразно ли включать то или иной фактор в уравнение регрессии, и в какой последовательности целесообразно водить факторы в модель множественной регрессии [2]. Алгоритм формирования множества возможных прогностических моделей представлен на рисунке 2.

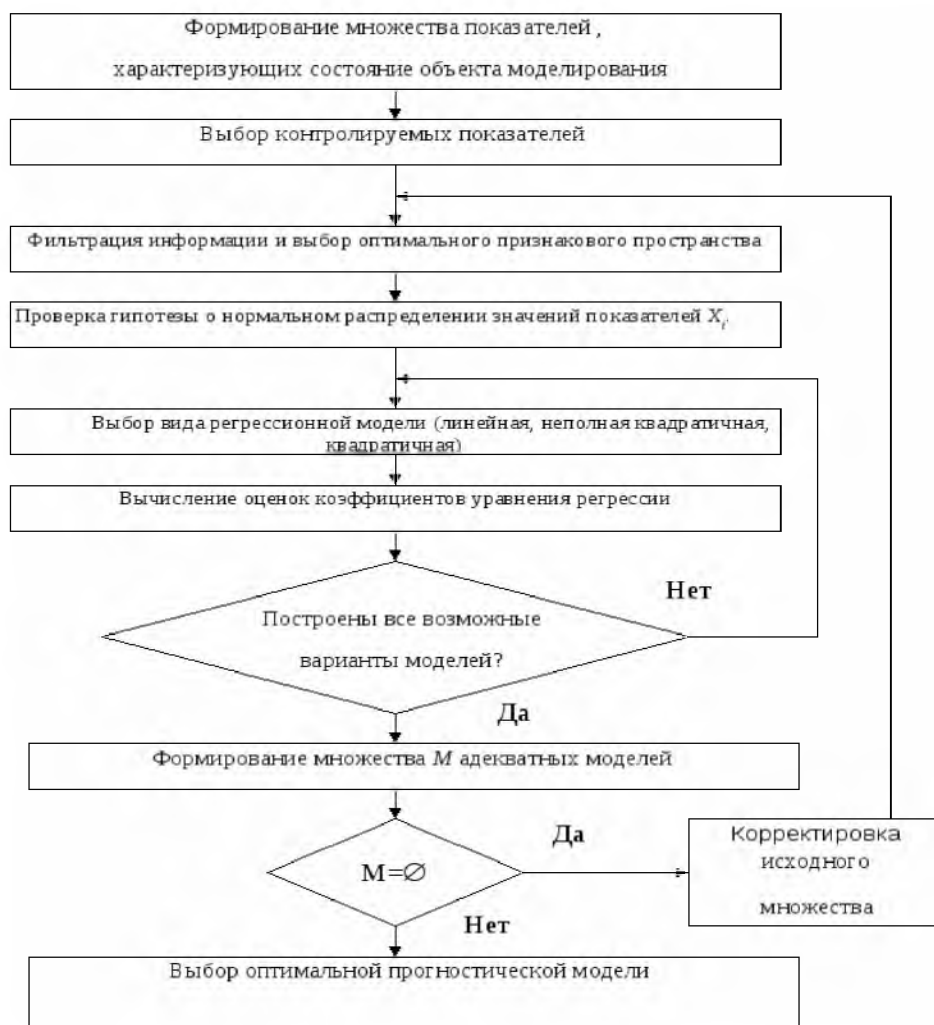


Рис.2. Схема алгоритма формирования прогностических моделей
Fig. 2. Chart of the formation of forecasting models

Общая структура программного комплекса следующая:

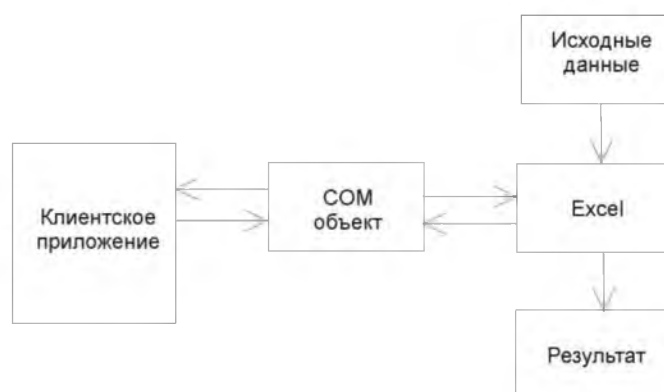


Рис.3. Общая структура программного комплекса
Fig.3. The overall structure of software

Для работы программы исходные данные экспортируются из MS Excel и отображаются на первой закладке программы [5]. Основная форма приложения с загруженными данными представлена ниже на рисунке.

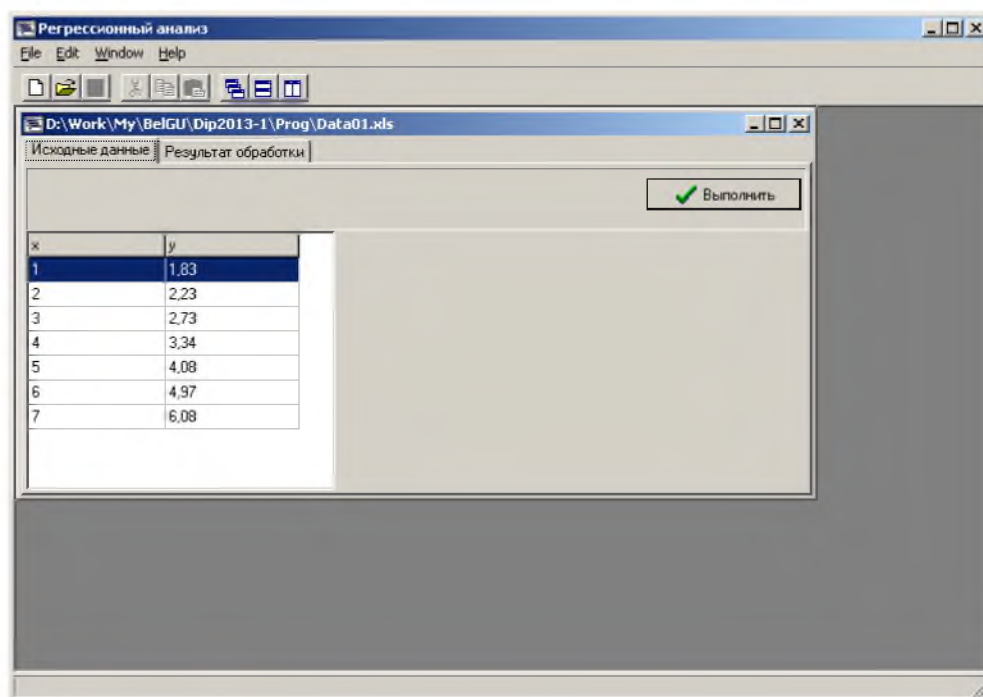


Рис.4. Основная форма MDI-приложения, с импортированными данными

Fig. 4. The main form of MDI-applications with imported data

На второй закладке выводятся результаты расчета по всем вариантам моделей регрессии, и осуществляется выбор оптимальной модели и выдача выводов и рекомендаций (рис 4):

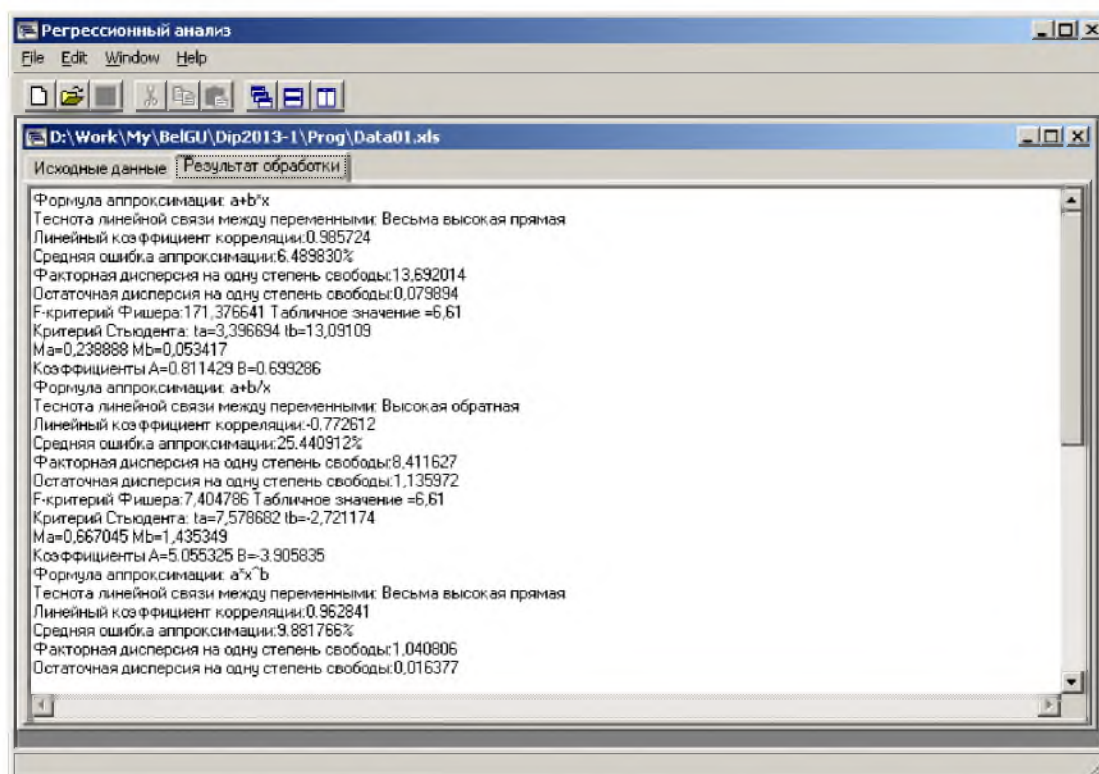


Рис. 5. Результаты расчета моделей парной регрессии

Fig. 5. The results of calculation of models of pair regression

После этого осуществляется запуск Excel через COM-объект и экспорт в Excel таблицы с расчетными данными. Затем через механизм COM-сервера клиентское приложение выполняет управление открытым экземпляром приложения Excel.Application. Следующим шагом создается новый лист с диаграммой для вывода расчетных значений (рис. 5):

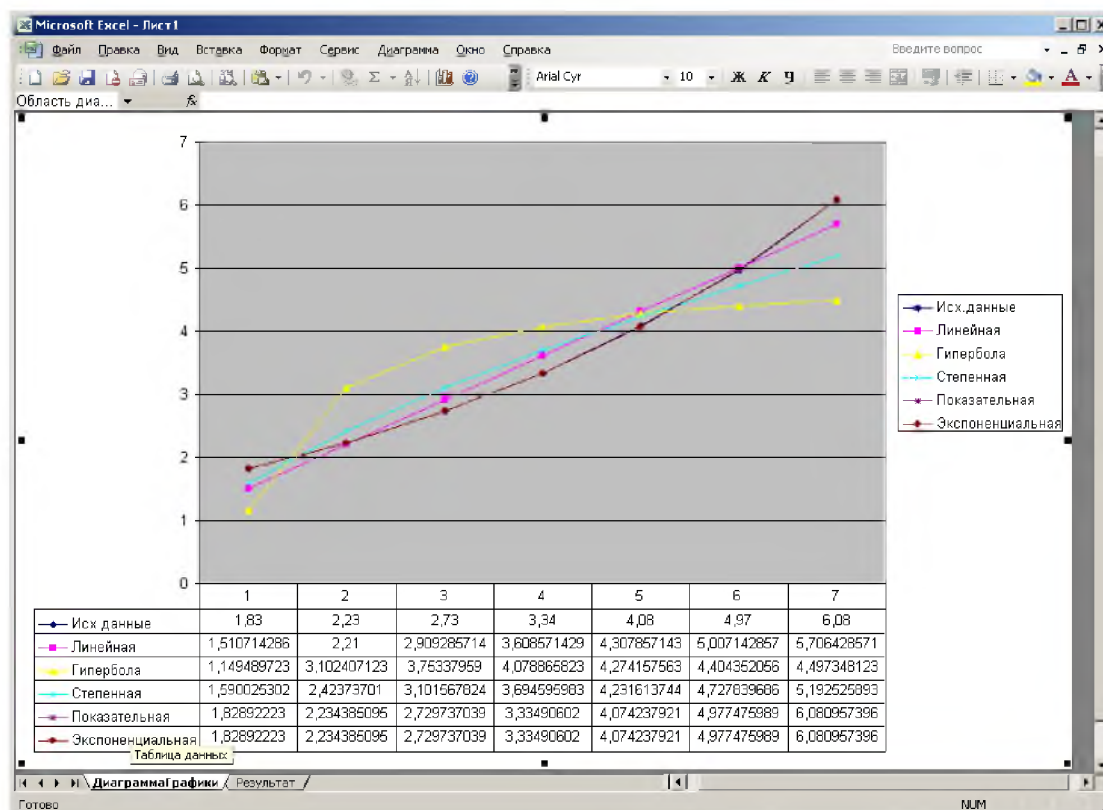


Рис. 6. Диаграмма вывода расчетных значений
Fig. 6. Output diagram of calculated values

Заключение

В данной статье был рассмотрен процесс разработки пакета компьютерной поддержки принятия решений для выбора корреляционно-регрессионных моделей. Разработанный пакет осуществляет построение основных моделей парной и множественной регрессии, выбор наиболее подходящей модели для описания эмпирических данных, оценка параметров, оценка статистической значимости модели и ее параметров, интерпретация результатов, выводы и пояснения системы.

Список литературы References

1. Путивцева Н.П., Игрунова С.В., Зайцева Т.В., Нестерова Е.В., Пусная О.П. Система поддержки принятия решений при выполнении проектов // Научные Ведомости БелГУ. Серия: Экономика. Информатика. - № 7 (204) 2015, выпуск 34/1.- С.170-174.
2. Путивцева Н.П., Игрунова С.В., Зайцева Т.В., Нестерова Е.В., Пусная О.П., Смородина Н.Н. Байесовская стратегия оценки достоверности выводов. Научные Ведомости БелГУ. Серия: История. Политология. Экономика. Информатика. - 2012 - №13(132), выпуск 23/1. - С. 180-183.
3. Зайцева Т.В., Нестерова Е.В., Игрунова С.В., Пусная О.П., Путивцева Н.П., Смородина Н.Н. Bajesovskaja strategija ocenki dostovernosti vyvodov. Nauchnye vedomosti BelGU. Ser. Istorija. Politologija. Jekonomika. Informatika. - 2012 - №13(132), vypusk 23/1. - S. 180-183.
4. Путивцева Н.П., Зайцева Т.В., Нестерова Е.В., Игрунова С.В., Гурьянова И.В. Эконометрика (базовый курс) НИУ «БелГУ» Учебное пособие Белгород: ООО «ГИК», 2013. – 203 с.
5. Путивцева Н.П., Зайцева Т.В., Нестерова Е.В., Игрунова С.В., Гурьянова И.В. Jekonometrika (bazovyj kurs) NIU «BelGU» Uchebnoe posobie Belgorod: ООО «ГИК», 2013. – 203 с.
6. Эконометрика Под редакцией И. И. Елисеевой, М.: Финансы и статистика, -2003 г.
7. Jekonometrika Pod redakciej I. I. Eliseevoj, M.: Finansy i statistika, -2003 g.
8. Delphi na primerah Pestrikov B., Masloboev A. SPb: BHV-Peterburg 2005 г., 496 с.
9. Delphi na primerah Pestrikov V., Masloboev A. SPb: BHV-Peterburg 2005 г., 496 с.